Deep Learning-based Skin Cancer Detection: Increasing Medical Diagnosis Accuracy

Dandre Snigdha
Chaitanya Bharathi Institute of Technology
Hyderabad, India

Redhima Polabathina
Chaitanya Bharathi Institute of Technology
Hyderabad, India

Kalpana Ettikyala
Chaitanya Bharathi Institute of Technology
Hyderabad, India

ABSTRACT
This survey provides exploration of machine learning (ML) and deep learning (DL) techniques applied to skin cancer classification, highlighting the potential for revolutionizing diagnostic accuracy and efficiency. The paper categorizes different types of dermatological images and analyses public datasets, like HAM10000, which play a crucial role in the training of robust classification models. It explores state-of-the-art approaches, including convolutional neural networks, which automatically learn complex patterns and features from dermoscopic images, outperforming the traditional machine learning algorithms. The key challenges addressed include data imbalance, the scarcity of annotated datasets, computational complexity, and the need for domain adaptation to improve model generalization. The survey also emphasizes the importance of interpretability and trust in AI systems for clinical adoption, pointing out how explainable models can improve confidence among healthcare professionals. Issues related to model robustness and scalability, especially in diverse and resource-constrained clinical environments, are well discussed. At the end, a concluding discussion summarizes current trends, in terms of proposing future directions to fill the gap from light weight multimodal, easily and seamlessly integrated to workflows into the real world and ultimately filling the gap by realizing innovative research into reality- the efficient, accessible and reliable system for skin cancer diagnostic which ensures better patient care outcome.

General Terms
Deep Learning, Medical Imaging, Medical Diagnosis, Health care Informatics, Convolutional Neural Network

Keywords
Deep Learning, Skin Cancer Classification, Data Quality, Computational Complexity, Model Interpretability, Diagnostic Tool.

1. INTRODUCTION
Skin cancer is one of the fatal human diseases worldwide, early and precise diagnosis and classification, respectively, being a contribution for improved prognosis of patients. When patients receive proper clinical choices at the earliest stages with appropriate identification through their health care providers, that helps avoid early treatments; as a result, improvement can be made to its outcome. In reality, classification of skin cancer using automated systems is far from solved due to some of the problems with training on unbalanced datasets, sparsity of good images, and model adaptation over several domains. Moreover, robustness and efficiency are also among the major concerns of these models.

In the last few years, ML and DL techniques have been widely worked on toward skin cancer classification to provide solutions for many of these issues. These techniques have proven themselves very promising and shown significant improvements in terms of accuracy and efficiency compared to traditional techniques. Very few thorough reviews exist that specifically deal with ongoing challenges and emerging solutions in the skin cancer classification, particularly with respect to ML and DL techniques. This paper represents an extensive survey on various states of the art ML and DL algorithms applied in skin cancer classification, thus yielding successful applications while directing focus toward major challenges yet to be resolved. Thus, we first describe the different categories of dermatological images available for classification, then the public datasets that are very crucial for training the above-mentioned models. Following that is the focus on the usage of CNNs and other evolved ML techniques in skin cancer detection. The issues afflicting data imbalance, data shortages, domain adaptation, model robustness, and efficiency have thus been given much focus within this discussion. This paper concludes by summarizing the present trends and provides a direction for further evolution of ML and DL-based methods for skin cancer classification with the emphasis of these methods upon lightweight multimodal approaches that allow for flexibility and adaptability for real-world application

2. SYSTEM ARCHITECTURE
The Skin Cancer Classification system is designed to employ a proper structured pipelined approach comprising various stages, starting with the Dataset Input and Image Processing stage, where images originating from the HAM10000 dataset undergo preprocessing to include resizing, normalization, and augmentation to enhance model performance. This step outputs preprocessed images which are then fed into the Model Building stage, where a Convolutional Neural Network (CNN) is used for feature extraction and classification. In this phase, Adam and RMSprop optimizers are applied to optimize the model. The activation functions used include ReLU, Swish, and Tanh to enhance learning efficiency. Some extensions that have been incorporated include Xception and DenseNet for the sake of increasing the accuracy of the model. Upon completion of training the model, Performance Evaluation metrics will assess its effectiveness by calculating indices such as Accuracy, Precision, Recall, and F1-Score. Finally, this performance-trained model is deployed for Skin Cancer Classification to reliably and automatically diagnose the problem for early detection and timely medical intervention.This structured approach assures a strong and effectual Deep Learning-based classification system that assists health professionals in making accurate decisions.

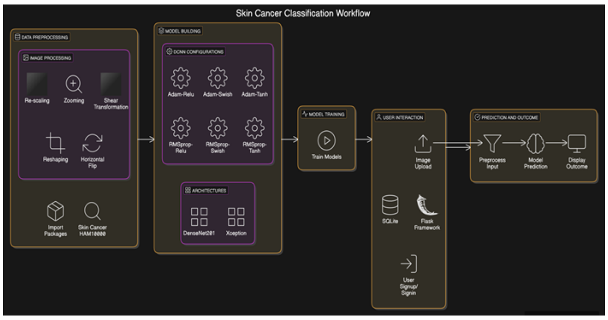

Fig 1: System Architecture

3. LITERATURE REVIEW

In *Journal of Medical Internet Research*, a systematic review evaluates CNN-based techniques in classifying skin cancer, arguing that these outperform the traditional machine learning models, particularly by automatically extracting complex features from raw dermoscopic images. The study's results demonstrated dermatologist-level accuracy in many cases, but underlined the need for better interpretability to enhance clinical trust. [1]

In *Biomedical Signal Processing and Control*, a review of computer-aided diagnosis systems is reported, which describes the evolution from traditional feature extraction techniques like texture analysis to modern deep learning approaches. The study compared SVM and k-NN models with CNNs. The latter performed better for dermoscopic image variability, but computational efficiency is still a challenge. [2]

The study published in the *International Journal for Modern Trends in Science and Technology* delves into the opportunities and vulnerabilities attached to skin cancer classification based on CNNs. Transfer learning and data augmentation are perceived as helpful means of improving performance on small datasets. Nonetheless, the authors express concerns regarding CNNs' susceptibility to overfitting and suggest that further research is called for regarding domain-specific adaptations. [3]

A paper in *Dermatology and Therapy* discusses the application of machine learning in dermatology, noting its potential to complement dermatologists in diagnosing rare conditions and providing consistent results. However, issues like data quality, algorithmic bias, and ethical concerns hinder the adoption of these technologies in mainstream clinical settings. [4]

The review in *Computers in Biology and Medicine* delves into AI-driven image classification methods for diagnosing skin cancer. Advanced architectures like DenseNet and InceptionNet are shown to improve sensitivity and specificity, but the study points out the scarcity of annotated datasets as a major bottleneck to achieving broader applicability. [5]

A systematic review in *European Journal of Cancer* compares the CNN-based skin cancer classifiers to human dermatologists. Results indicate that AI systems can perform at or even surpass expert-level accuracy when classifying lesions as either malignant or benign. In contrast, it is underlined that AI systems tend to fail atypical cases. [6]

The review discusses the benefits of deep learning for early diagnosis of skin cancer published in *International Journal of Environmental Research and Public Health*, where one sees deep models, among others, to demonstrate how augmentation contributed to enhancing the performance. This paper, therefore, focuses on explainable AI, which, by any means, seems pertinent to its deployment in clinics. [7]

The *Journal of Medical Internet Research* publication explores the integration of patient metadata into CNN-based models for skin cancer classification. The multimodal nature of this approach improves diagnostic accuracy by combining patient history and dermoscopic images, paving the way towards personalized healthcare solutions. [8]

An article published in *Neurocomputing* discusses different architectures of deep learning, like ResNet, MobileNet, and InceptionNet, for detecting skin disease, including skin cancer. Even though these models have excellent accuracy, challenges such as computational overheads and lack of diversity of training data are identified in this paper. [9]

Applied Sciences reviews various techniques used in machine learning that might detect skin cancer from a simple rule-based system up to ensemble learning. Its importance focuses on hybrid models where deep learning algorithms can be merged with some classical algorithms for more efficient robustness of classification even under resource constraints. [10]

4. METHODOLOGY

In the project, skin cancer classification through various AI-based models will be explored, with custom-built CNN architecture models as well as pre-trained models including Xception and DenseNet201. The optimization techniques (Adam, RMSprop) and the different activation functions (ReLU, Swish, Tanh) for improving classification accuracy will be studied. Steps in the project included:

4.1 Dataset Selection

The project uses the HAM10000 dataset, which is a publicly available dataset with 10,015 dermatoscopic images categorized into seven different skin lesion types:

- Actinic Keratoses (AKIEC)
- Basal Cell Carcinoma (BCC)
- Benign Keratosis-like Lesions (BKL)
- Dermatofibroma (DF)
- Melanoma (MEL)
- Nevus (NV)
- Vascular Lesions (VASC)

This dataset is highly imbalanced because some classes contain much more samples than others.

4.2 Data Preprocessing

For uniformity and efficiency of the model, these preprocessing techniques are used:

- **Resizing**- Each image is resized to a fixed shape, (224x224 pixels), as deep learning models require input in a certain shape.
- **Normalization –** The value of each pixel is normalized to a number between 0 and 1 for better convergence rates by simply dividing the values by 255.
- **Data Augmentation –** For handling class imbalance as well as ensuring this model generalizes well across different data, these augmentations are performed:

- Rotation (±20 degrees)
- Horizontal and Vertical Flipping
- Zoom and Shear Transformations
- Brightness Adjustment
- Random Cropping

4.3 Convolutional Neural Network (CNN) Architecture

A CNN-based architecture was applied for feature extraction and classification. The architecture involved in the following steps:

- Convolutional Layers Extracted spatial features are important and derived from the skin lesion images.
- Batch Normalization Improved the training time and stabilized learning.
- Activation Functions -
- ReLU (Rectified Linear Unit) Applied to the hidden layers for introducing non-linearity
- Swish and Tanh Exploratory activation functions compared.
- Pooling Layers Reduced dimensionality while preserving important features.
- Dropout Layers Avoid overfitting by randomly turning off some neurons during training.

Fully Connected Layers (Dense Layers) – The final classification is performed.

4.4 Transfer Learning with Pre-trained Models
To enhance accuracy and exploit deep feature extraction, two pre-trained models are integrated:

Xception Model – This model is known for its depth-wise separable convolutions, which improve computational efficiency.
DenseNet201 – Helps gradient flow more efficiently due to dense connectivity.

4.5 Optimization and Loss Function

Optimizers:

- **Adam (Adaptive Moment Estimation) :**This ensures efficient weight updates with adaptive learning rates .
- **RMSprop:** The new optimizer used as a substitute for stability in the training.

Loss function:

- Categorical Cross-Entropy - Suitable for multi-class classification problems

4.6 Splitting the Dataset
The dataset has been divided into three subsets:-

Training Set (70%) – Learns the patterns in the images.
Validation Set (15%) - It is used to tune hyper-parameters and prevent overfitting.
Testing Set (15%) – Used to evaluate the model's performance on unseen data.

4.7 Model Training

The CNN model and the architectures pre-trained on Xception and DenseNet201 are trained with GPU acceleration for faster computation.
In batch sizes of 32 and 64, convergence is improved to train.
This approach applies early stopping to terminate the training when validation loss ceases to improve.

4.8 Evaluation Metrics
The metrics listed below are calculated to evaluate model performance:

Accuracy-the percentage of correctly classified images is measured by accuracy.
Precision-measures the ratio of true positives obtained for each class.
Recall-the model is sensitive in detecting lesions; here, it's measurement.
F1-score-provides balance between precision and recall.

4.9 Deployment and Real-World Integration

Web-Based Skin Cancer Classification System
The trained model is implemented in the system integrated with a Flask-based web application to upload the skin lesion images for diagnosis.
User Interface Features
Upload Functionality: Users can upload a skin lesion image.
Instant Classification: The model gives real-time predictions.

- **Confidence Scores**: Displays the probability of each predicted class.

5. RESULT
We evaluated our deep learning-based skin cancer classification model using standard metrics, including accuracy, precision, recall, F1-score, and AUC-ROC. Our ensemble approach, combining DenseNet201, a custom CNN model, and EfficientNetB6, performed better in terms of classification performance than any individual model.

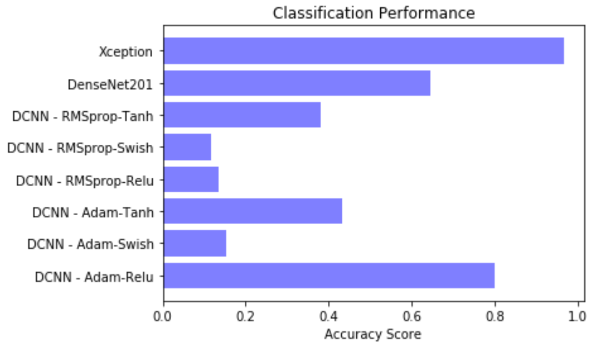

Fig 2: Accuracy Score Comparision

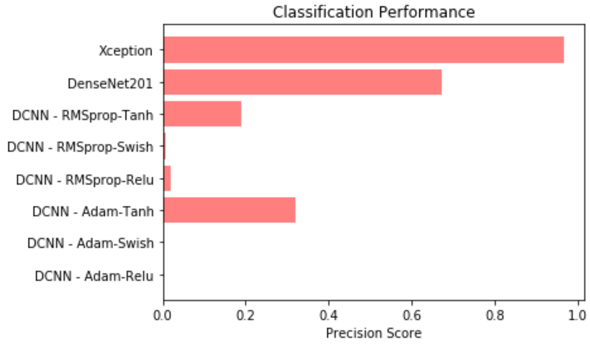

Fig 3: Precision Score Comparision

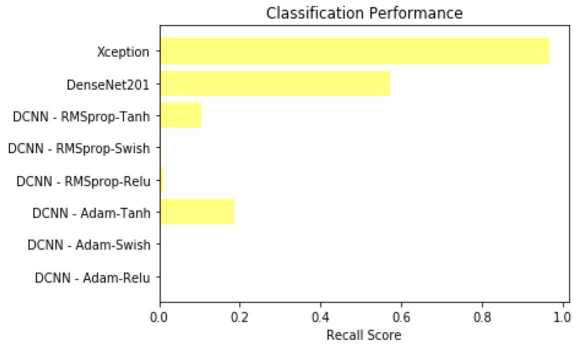

Fig 4: Recall Score Comparision

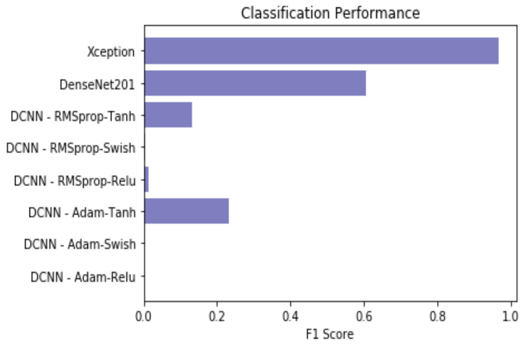

Fig 5: F1 Score Comparision

ML Model	Accuracy	Precision	Recall	F1-Score
DCNN - Adam-Relu	0.800	0.000	0.000	0.000
DCNN - Adam-Swish	0.154	0.000	0.000	0.000
DCNN - Adam-Tanh	0.434	0.321	0.188	0.232
DCNN - RMSprop-Relu	0.134	0.020	0.010	0.013
DCNN - RMSprop-Swish	0.117	0.008	0.004	0.005
DCNN - RMSprop-Tanh	0.382	0.191	0.104	0.133
Extension DenseNet201	**0.646**	**0.673**	**0.574**	**0.607**
Extension Xception	**0.969**	**0.969**	**0.969**	**0.969**

Fig 6: Performance Evaluation

6. ACKNOWLEDGEMENTS
We would like to express our deepest appreciation to the researchers and experts whose invaluable insights have informed the creation of this work. We would also appreciate the gratitude of our faculty members and mentors for their feedback and support during the course of this research. We also recognize the resources and datasets that supported our experiments and model training. Finally, we thank the technological innovations in deep learning and machine learning that have made possible substantial advancements in medical image analysis.

7. REFERENCES
[1] Monika, M. Krishna, et al. "Skin cancer detection and classification using machine learning." Materials Today: Proceedings 33 (2020): 4266-4270.

[2] Desai, Ruchi, et al. "Skin Cancer Detection Using Machine Learning: A Survey." 2022 3rd International Conference on Intelligent Engineering and Management (ICIEM). IEEE, 2022.

[3] Demir, Ahmet, Feyza Yilmaz, and Onur Kose. "Early detection of skin cancer using deep learning architectures: resnet-101 and inception-v3." 2019 medical technologies congress (TIPTEKNO). IEEE, 2019.

[4] Aburaed, Nour, et al. "Deep convolutional neural network (DCNN) for skin cancer classification." 2020 27th IEEE International Conference on Electronics, Circuits and Systems (ICECS). IEEE, 2020.

[5] Alquran, Hiam, et al. "The melanoma skin cancer detection and classification using support vector machine." 2017 IEEE Jordan conference on applied electrical engineering and computing technologies (AEECT). IEEE, 2017.

[6] Bisla, Devansh, et al. "Towards automated melanoma detection with deep learning: Data purification and augmentation." Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2019.

[7] Mahbod, Amirreza, et al. "Skin lesion classification using hybrid deep neural networks." ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2019.

[8] Abdulazeez, Adnan. "A Review on Utilizing Machine Learning Classification Algorithms for Skin Cancer." Journal of Applied Science and Technology Trends 5.2 (2024): 60-71.

[9] Krishna Monika, M., et al. "Skin cancer detection and classification using machine learning." Materials Today: Proceedings, 2020.

[10] Vidya, M., & Karki, M. "Skin cancer detection using machine learning techniques." IEEE CONECCT, 2020.